

ABSTRACT OF THE DISCLOSURE

A method for creating a structured document, wherein a structured document comprises a plurality of content elements wrapped in pairs of tags, includes parsing a document of a particular type containing content into a plurality of content elements; and for
5 each content element, suggesting an optimal tag according to a tag suggestion procedure. The tag suggestion procedure includes providing sample data which has been converted into a structured sample document; deriving a set of tags from the structured sample document; evaluating the set of tags according to tag suggestion criteria to determine an optimal tag for
10 the content element. The optimal tag may be a single tag or a pattern of tags which maximizes a similarity function with patterns found in the sample data.